



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/53, 15/81, 9/02, 5/10, 1/19, C12P 7/64, C11B 1/00, A61K 31/20, A23L 1/30		A1	(11) International Publication Number: WO 98/46763 (43) International Publication Date: 22 October 1998 (22.10.98)
<p>(21) International Application Number: PCT/US98/07126</p> <p>(22) International Filing Date: 10 April 1998 (10.04.98)</p> <p>(30) Priority Data: US 08/834,655 11 April 1997 (11.04.97) US</p> <p>(63) Related by Continuation (CON) or Continuation-in-Part (CIP) to Earlier Application US 08/834,655 (CIP) Filed on 11 April 1997 (11.04.97)</p> <p>(71) Applicants (for all designated States except US): CALGENE LLC [US/US]; 1920 Fifth Street, Davis, CA 95616 (US). ABBOTT LABORATORIES [US/US]; 100 Abbott Park Road, Abbott Park, IL 60064-3500 (US).</p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (for US only): KNUTZON, Deborah [US/US]; 6110 Rockhurst Way, Granite Bay, CA 95746 (US). MUKERJI, Pradip [US/US]; 1069 Arcaro Drive, Gahanna, OH 43230 (US). HUANG, Yung-Sheng [CA/US]; 2462 Danvers Court, Upper Arlington, OH 43220 (US). THURMOND, Jennifer [US/US]; 3702 Adirondack, Colum-</p>		<p>bus, OH 43231 (US). CHAUDHARY, Sunita [IN/US]; 3419 Woodbine Place, Pearland, TX 77584 (US). LEONARD, Amanda, Eun-Yeong [US/US]; 581 Shadewood Court, Gahanna, OH 43230 (US).</p> <p>(74) Agents: WARD, Michael, R. et al.; Limbach & Limbach L.L.P., 2001 Ferry Building, San Francisco, CA 94111-4262 (US).</p> <p>(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p>Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</p>	
<p>(54) Title: METHODS AND COMPOSITIONS FOR SYNTHESIS OF LONG CHAIN POLYUNSATURATED FATTY ACIDS</p> <p>(57) Abstract</p> <p>The present invention relates to fatty acid desaturases able to catalyze the conversion of oleic acid to linoleic acid, linoleic acid to γ-linolenic acid, or of alpha-linolenic acid to stearidonic acid. Nucleic acid sequences encoding desaturases, nucleic acid sequences which hybridize thereto, DNA constructs comprising a desaturase gene, and recombinant host microorganism or animal expressing increased levels of a desaturase are described. Methods for desaturating a fatty acid and for producing a desaturated fatty acid by expressing increased levels of a desaturase are disclosed. Fatty acids, and oils containing them, which have been desaturated by a desaturase produced by recombinant host microorganisms or animals are provided. Pharmaceutical compositions, infant formulas or dietary supplements containing fatty acids which have been desaturated by a desaturase produced by a recombinant host microorganism or animal also are described.</p>			
<pre> graph TD PA[Palmitic Acid (C16)] --> SA[Stearic Acid (C18)] SA --> OA[18:1 Δ9] OA --> LA[18:2 Δ9, 12] LA --> GLA[α-18:3 Δ9, 12, 15] GLA --> SDA[Stearidonic] SDA --> DHGLA[DH-γ-Linolenic] DHGLA --> 20:3[20:3] 20:3 --> 20:4[20:4] 20:4 --> AA[Arachidonic] AA --> PG1[Prostaglandin G1] AA --> PG2[Prostaglandin G2] AA --> PGF2a[Prostaglandin E2, F2a] SA --> P1[Palmitoleic (C16Δ9)] P1 --> P2[Higher saturated fatty acid] OA --> P3[18:1 Δ9] P3 --> P4[18:2 Δ9, 12] P4 --> P5[18:3 Δ9, 12, 15] P5 --> P6[18:4 Δ9, 12, 15] P6 --> P7[Stearidonic] P7 --> P8[DH-γ-Linolenic] P8 --> P9[20:3] P9 --> P10[20:4] P10 --> P11[Arachidonic] P11 --> P12[Prostaglandin G1] P11 --> P13[Prostaglandin G2] P11 --> P14[Prostaglandin E2, F2a] </pre>			

The present invention is further directed to transgenic oils in pharmaceutically acceptable carriers. The present invention is further directed to nutritional supplements, cosmetic agents and infant formulae containing transgenic oils.

5 The present invention is further directed to a method for obtaining altered long chain polyunsaturated fatty acid biosynthesis comprising the steps of: growing a microbe having cells which contain a transgene which encodes a transgene expression product which desaturates a fatty acid molecule at carbon 6 or 12 from the carboxyl end of said fatty acid molecule, wherein the transgene
10 is operably associated with an expression control sequence, under conditions whereby the transgene is expressed, whereby long chain polyunsaturated fatty acid biosynthesis in the cells is altered.

15 The present invention is further directed toward pharmaceutical compositions comprising at least one nutrient selected from the group consisting of a vitamin, a mineral, a carbohydrate, a sugar, an amino acid, a free fatty acid, a phospholipid, an antioxidant, and a phenolic compound.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows possible pathways for the synthesis of arachidonic acid (20:4 Δ5, 8, 11, 14) and stearidonic acid (18:4 Δ6, 9, 12, 15) from palmitic acid (C₁₆) from a variety of organisms, including algae, *Mortierella* and humans. These PUFAs can serve as precursors to other molecules important for humans and other animals, including prostacyclins, leukotrienes, and prostaglandins, some of which are shown.

25 Figure 2 shows possible pathways for production of PUFAs in addition to ARA, including EPA and DHA, again compiled from a variety of organisms.

Figure 3A-E shows the DNA sequence of the *Mortierella alpina* Δ6-desaturase and the deduced amino acid sequence:

Figure 3A-E (SEQ ID NO 1 Δ6 DESATURASE cDNA)

Figure 3A-E (SEQ ID NO 2 Δ 6 DESATURASE AMINO ACID)

Figure 4 shows an alignment of a portion of the *Mortierella alpina* Δ 6-desaturase amino acid sequence with other related sequences.

5 Figure 5A-D shows the DNA sequence of the *Mortierella alpina* Δ 12-desaturase and the deduced amino acid sequence:

Figure 5A-D (SEQ ID NO 3 Δ 12 DESATURASE cDNA)

Figure 5A-D (SEQ ID NO 4 Δ 12 DESATURASE AMINO ACID).

Figures 6A and 6B show the effect of different expression constructs on expression of GLA in yeast.

10 Figures 7A and 7B show the effect of host strain on GLA production.

Figures 8A and 8B show the effect of temperature on GLA production in *S. cerevisiae* strain SC334.

Figure 9 shows alignments of the protein sequence of the Ma 29 and contig 253538a.

15 Figure 10 shows alignments of the protein sequence of Ma 524 and contig 253538a.

BRIEF DESCRIPTION OF THE SEQUENCE LISTINGS

SEQ ID NO:1 shows the DNA sequence of the *Mortierella alpina* Δ 6-desaturase.

20 SEQ ID NO:2 shows the protein sequence of the *Mortierella alpina* Δ 6-desaturase.

SEQ ID NO:3 shows the DNA sequence of the *Mortierella alpina* Δ 12-desaturase.

25 SEQ ID NO:4 shows the protein sequence of the *Mortierella alpina* Δ 12-desaturase.

SEQ ID NO:5-11 show various desaturase sequences.

SEQ ID NO:13-18 show various PCR primer sequences.

SEQ ID NO:19 and SEQ ID NO:20 show the nucleotide and amino acid sequence of a *Dictyostelium discoideum* desaturase.

5 SEQ ID NO:21 and SEQ ID NO:22 show the nucleotide and amino acid sequence of a *Phaeodactylum tricornutum* desaturase.

SEQ ID NO:23-26 show the nucleotide and deduced amino acid sequence of a *Schizochytrium* cDNA clone.

SEQ ID NO: 27-33 show nucleotide sequences for human desaturases.

10 SEQ ID NO:34 - SEQ ID NO:40 show peptide sequences for human desaturases.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

In order to ensure a complete understanding of the invention, the following definitions are provided:

15 **Δ5-Desaturase:** Δ5 desaturase is an enzyme which introduces a double bond between carbons 5 and 6 from the carboxyl end of a fatty acid molecule.

Δ6-Desaturase: Δ6-desaturase is an enzyme which introduces a double bond between carbons 6 and 7 from the carboxyl end of a fatty acid molecule.

Δ9-Desaturase: Δ9-desaturase is an enzyme which introduces a double bond between carbons 9 and 10 from the carboxyl end of a fatty acid molecule.

20 **Δ12-Desaturase:** Δ12-desaturase is an enzyme which introduces a double bond between carbons 12 and 13 from the carboxyl end of a fatty acid molecule.

Fatty Acids: Fatty acids are a class of compounds containing a long hydrocarbon chain and a terminal carboxylate group. Fatty acids include the 25 following:

Fatty Acid		
12:0	lauric acid	
16:0	palmitic acid	

FIG. 3A

CGACACTCCT TCCTCTCTCT CACCCGTCCCT AGTCCCCCTC AACCCCCCTC TTTGACAAAG
 60 *
 ACAACAAACC ATG GCT GCT GCT CCC AGT GTG AGG ACG TTT ACT CGG GCC GAG
 Met Ala Ala Pro Ser Val Arg Thr Phe Thr Arg Ala Glu *.../*
 120

GTT TTG AAT GCC GAG GCT CTG AAT GAG GGC AAG AAG GAT GCC GAG GCA
 Val Leu Asn Ala Glu Ala Leu Asn Glu Gly Lys Asp Ala Glu Ala *.../*
 180

CCC TTG ATG ATC ATC GAC AAC AAG GTG TAC GAT GTC CGC GAG TTG
 Pro Phe Leu Met Ile Ile Asp Asn Lys Val Tyr Asp Val Arg Glu Phe *.../*
 240

GTC CCT GAT CAT CCC GGT GGA AGT GTG ATT CTC ACG CAC GTT GGC AAG
 Val Pro Asp His Pro Gly Gly Ser Val Ile Leu Thr His Val Gly Lys *.../*
 300

GAC GGC ACT GAC GTC TTT GAC ACT TTT CAC CCC GAG GCT GCT TGG GAG
 Asp Gly Thr Asp Val Phe Asp Thr Phe His Pro Glu Ala Ala Trp Glu
 360

ACT CTT GCC AAC TTT TAC GCT GGT GAT ATT GAC GAG AGC GAC CGC GAT
 Thr Leu Ala Asn Phe Tyr Val Gly Asp Ile Asp Glu Ser Asp Arg Asp *.../*
 420

ATC AAG AAT GAT GAC TTT GCG GCG GAG GTC CGC AAG CTG CGT ACC TTG
 Ile Lys Asn Asp Asp Phe Ala Ala Glu Val Arg Lys Leu Arg Thr Leu *.../*

FIG. 3B

420 *
 TTC CAG TCT CTT GGT TAC TAC GAT TCT TCC AAG GCA TAC TAC GCC TTC
 Phe Gln Ser Leu Gly Tyr Tyr Asp Ser Ser Lys Ala Tyr Tyr Ala Phe 166

480 *
 AAG GTC TCG TTC AAC CTC TGC ATC TGG GGT TTG TCG ACG GTC ATT GTG
 Lys Val Ser Phe Asn Leu Cys Ile Trp Gly Leu Ser Thr Val Ile Val 222

540 *
 GCC AAG TCG GGC CAG ACC TCG ACC CTC GCC AAC GTG CTC TCG GCT GCG
 Ala Lys Trp Gly Gln Thr Ser Thr Leu Ala Asn Val Leu Ser Ala Ala 188

600 *
 CTT TGG GCT CTG TGC TGG CAG CAG TGC GGA TGG TTG GCT CAC GAC TTT
 Leu Leu Gly Leu Phe Trp Gln Gln Cys Gly Trp Leu Ala His Asp Phe 254

660 *
 TTG CAT CAC CAG GTC TTC CAG GAC CGT TTC TGG GGT GAT CTT TTC GGC
 Leu His His Gln Val Phe Gln Asp Arg Phe Trp Gly Asp Leu Phe Gly 290

720 *
 GCC TTC TTG GGA GGT GTC TGC CAG GGC TTC TCG TCC TCG TGG TGG AAG
 Ala Phe Leu Gly Val Cys Gln Gly Phe Ser Ser Trp Trp Lys 216

780 *
 GAC AAG CAC AAC ACT CAC CAC GCC GCC CCC AAC GTC CAC GGC GAG GAT
 Asp Lys His Asn Thr His Ala Ala Pro Asn Val His Gly Glu Asp 274

FIG. 3C

CCC GAC ATT GAC ACC CAC CCT CTG TRG ACC TGG AGT GAG CAT GCG TTG
 Pro Asp Ile Asp Thr His Pro Leu Leu Thr Trp Ser Glu His Ala Leu

GAG ATG TTC TCG GAT GTC CCA GAT GAG GAG CTG ACC CGC ATG TGG TCG
 Glu Met Phe Ser Asp Val Pro Asp Glu Glu Leu Thr Arg Met Trp Ser

840 *

CGT TTC ATG GTC CTG AAC CAG ACC TGG TTT TAC TTC CCC ATT CTC TCG TCG
 Arg Phe Met Val Leu Asn Gln Thr Trp Phe Tyr Phe Pro Ile Leu Ser

900

TTT GCC CGT CTC TCC TGG TGC CTC CAG TCC ATT CTC TTT GTG CTG CCT
 Phe Ala Arg Leu Ser Trp Cys Leu Gln Ser Ile Leu Phe Val Leu Pro

960 *

AAC GGT CAG GCC CAC AAG CCC TCG GGC GCG CGT GTG CCC ATC TCG TCG TCG
 Asn Gly Gln Ala His Lys Pro Ser Gly Ala Arg Val Pro Ile Ser Leu

1020

GTC GAG CAG CTG TCG CTT GCG ATG CAC TGG ACC TGG TAC CTC GCC ACC
 Val Glu Gln Leu Ser Leu Ala Met His Trp Thr Trp Tyr Leu Ala Thr

ATG TTC CTG TTC ATC AAG GAT CCC GTC AAC ATG CTG GTG TAC TTT TTG
 Met Phe Leu Phe Ile Lys Asp Pro Val Asp Met Leu Val Tyr Phe Leu

1080 *

GTG TCG CAG GCG GTG TGC GGA AAC TTT GCG ATC GTG TTC TCG CTC
 Val Ser Gln Ala Val Cys Gly Asn Leu Leu Ala Ile Val Phe Ser Leu

FIG. 3D

AAC CAC AAC GGT ATG CCT GTG ATC TCG AAG GAG GCG GTC GAT ATG
 Asn His Asn Gly Met Pro Val Ile Ser Lys Glu Ala Val Asp Met 1140

GAT TTC TTC ACG AAG CAG ATC ATC ACG GGT CGT GAT GTC CAC CCG GGT
 Asp Phe Phe Thr Lys Gln Ile Ile Thr Gly Arg Asp Val His Pro Gly 1200

CTA TTT GCC AAC TGG TTC ACG GGT GGA TTG AAC TAT CAG ATC GAG CAC
 Leu Phe Ala Asn Trp Phe Thr Gly Gly Leu Asn Tyr Gln Ile Glu His 1260

CAC TTG TTC CCT TCG ATG CCT CGC CAC AAC TTT TCA AAG ATC CAG CCT
 His Leu Phe Pro Ser Met Pro Arg His Asn Phe Ser Lys Ile Gln Pro 1320

GCT GTC GAG ACC CTG TGC AAA AAG TAC AAT GTC CGA TAC }CAC ACC ACC
 Ala Val Glu Thr Leu Cys Lys Tyr Asn Val Arg Tyr }His Thr Thr 1380

GGT ATG ATC GAG GGA ACT GCA GAG GTC TTT AGC CGT CTG AAC GAG GTC
 Gly Met Ile Glu Gly Thr Ala Glu Val Phe Ser Arg Leu Asn Glu Val 1440

TCC AAG GCT GCC TCC AAG ATG GGT AAG GCG CAG TAAAAAAA AAAAGGAC
 Ser Lys Ala Ala Ser Lys Met Gly Lys Ala Gln

FIG. 3E

1500 * GTTTTTTC GCCAGTGCC GTGCCTGTGC CTGCTTCCT TGTCAAGTCG AGCGTTCTG
1560 * GAAAGGATCG TTCAGTGCAG TATCATCATT CTCCCTTTAC CCCCGCTCA TATCTCATTC
ATTCTCTTA TAAACAACT TGTTCCCCCC TTCAACCG

FIG. 5A

GTCCCCCTGTC GCTGGGCA CACCCCATCC TACCTGGTC CCTCTGGTT TCTCTGGCC
 60 *
 CCACCGGTCTC TTCTCCACCC TCCGAGACGA CTGCAACTGT ATTAGGAAAC CGRCAAATAC
 120 *
 ACGATTCTT TTAACTCTGC ACCAACTCAA AATCCTAAC CGGAACCCCTT TTTCAGG ATG
 180 *
 GCA CCT CCC AAC ACT ATC GAT GCC GGT TTG ACC CAG CTC CAT ATC AGC
 Ala Pro Pro Asn Thr Ile Asp Ala Gly Leu Thr Gln Arg His Ile Ser
 240
 ACC TCG GCC CCA AAC TCG GCC AAG CCT CCC TTG GAG CGC CGC AAC TAC CAG
 Thr Ser Ala Pro Asn Ser Ala Lys Pro Ala Phe Glu Arg Asn Tyr Gln
 300 *
 CTC CCC GAG TTC ACC ATC AAG GAG ATC CGA GAG TGC ATC CCT GCC CAC
 Leu Pro Glu Phe Thr Ile Lys Glu Ile Arg Glu Cys Ile Pro Ala His
 360 *
 TGC TTT GAG CGC TCC GGT CTC CGT CGT CTC TTG CAC GTC GCC ATC GAT
 Cys Phe Glu Arg Ser Gly Leu Arg Gly Leu Cys His Val Ala Ile Asp
 420 *
 CTG ACT TGA GCG TCG CTC TTG TTG TGC GCT GCG ACC CAG ATC GAC AAG
 Leu Thr Trp Ala Ser Leu Leu Phe Leu Ala Ala Thr Gln Ile Asp Lys
 480 *
 TTT GAG AAT CCC TTG ATC CGC TAT TTG GCC TGG CCT GTC TAC TGG ATC
 Phe Glu Asn Pro Leu Ile Arg Tyr Leu Ala Trp Pro Val Tyr Trp Ile

FIG. 5B

480 ATG CAG GGT ATT GTC TGC ACC GGT GTC TGG CTC GCT CAC GAG TGT
 Met Glu Gly Ile Val Cys Thr Gly Val Trp Val Leu Ala His Glu Cys
 540 GGT CAT CAG TCC TCG ACC TCC AAG ACC CTC AAC AAC ACA GTT GGT
 Gly His Glu Ser Phe Ser Thr Ser Lys Thr Leu Asn Asn Thr Val Gly
 600 TCG ATC TGG CAC TCG ATT GTC CCC TAC CTC TGG AGA ATC
 Trp Ile Leu His Ser Met Leu Leu Val Pro Tyr His Ser Thr Arg Ile
 660 TCG CAC TCG AAG CAC CAC AAG GCC ACT GCC CAT ATG ACC AAG GAC CAG
 Ser His Ser Lys His His Lys Ala Thr Gly His Met Thr Lys Asp Glu
 720 GTC TTT GTC CCC AAG ACC CGC TCC CAG GTC GCC TTC CCT CCC AAG GAG
 Val Phe Val Pro Lys Thr Arg Ser Glu Val Gly Leu Pro Pro Lys Glu
 780 AAC GCT GCT GCT GCG GTC ATT GTC TGG CTC TGG ATG TCC GTC CAC CTC GAT
 Asn Ala Ala Ala Val Glu Glu Asp Met Ser Val Val Ile Glu Asn Leu Asp
 840 TTC CGA TGG CCC GCG TAC CTG ATT ATG AAC GCC TCT GGC CAA GAC TAC
 Phe Gly Trp Pro Ala Tyr Leu Ile Asn Ala Ser Gly Glu Asp Tyr

FIG. 5C

900 *
 CGC CGC TGG ACC TCC CAC TTC CAC ACG TAC TCG CCC ATC TTT GAG CCC
 Gly Arg Trp Thr Ser His Phe His Thr Tyr Ser Pro Ile Phe Glu Pro

 CGC AAC TTT TTC GAC ATT ATT ATC TCG GAC CTC GCT GTC TTG GCT GGC
 Arg Asn Phe Phe Asp Ile Ile Ile Ser Asp Leu Gly Val Leu Ala Ala
 960 *

 CTC GGT GCC CTG ATC TAT GCC TCC ATG CAG TTG CTC TCG ACC GTC
 Leu Gly Ala Leu Ile Tyr Ala Ser Met Gln Leu Ser Leu Leu Thr Val
 1020 *

 ACC AAG TAC TAT ATT GTC CCC TAC CTC TTT GTC AAC TTT TTG TTG GTC
 Thr Lys Tyr Tyr Ile Val Pro Tyr Leu Phe Val Asp Phe Trp Leu Val

 1080 *

 CTG ATC ACC TTC TTG CAG CAC ACC GAT CCC AAG CTG CCC CAT TAC TAC CGC
 Leu Ile Thr Phe Leu Gln His Thr Asp Pro Lys Leu Pro His Tyr Arg
 1140 *

 GAG GGT GCC TGG AAT TTC CAG CCT GGA CCT CTT TGC ACC GTT GAC CGC
 Glu Gly Ala Thr Asp Phe Gln Arg Gly Ala Leu Cys Thr Val Asp Arg

 1200 *

 TCG TTT GGC AAG TTC TTG GAC CAT ATT GTC CAC GGC ATT GTC CAC ACC
 Ser Phe Gly Lys Phe Leu Asp His Met Phe His Gly Ile Val His Thr

 CAT GNG GCC CAT CAC TTG TGC TGG CAA ATG CCG TTG TAC CAT GCT GAG
 His Val Ala His His Leu Phe Ser Gln Met Pro Phe Tyr His Ala Glu

FIG. 5D

GAA CCT ACC TAT CAT CTC AAG AAA CTG CTG GGA GAG TAC TAT GTG TAC Glu Ala Thr Tyr His Leu Lys Lys Leu Leu Gly Glu Tyr Tyr Val Val Tyr	1260	GAC CCA TCC CCG ATC GTC GTC GTC GTC TGG AGG TCG TTC CGT GAG TGC Asp Pro Ser Pro Ile Val Val Ala Val Val Val Val Phe Arg Ser Phe Arg Glu Cys	1320	CGA TTC GTG GAG GAT CAG GGA GAC GAC GTC GTC TCT TCT AAG AAG TAAAGA Arg Phe Val Glu Asp Gln Gln Gly Asp Val Val Phe Phe Lys Lys	1380	AAAAGACAAAT GGACCACACA CAACCTTGTCT TCTACAGACC TACGTATCAT GTAGCCATAC CACTTCTATAA AAGAACATGAA GCTCTAGAGG CGGTGTCATTC GCGCCCTTC	1440
--	------	--	------	---	------	---	------

FastA Match of ma29 and contig 253538a

SCORES Init1: 117 Initn: 225 Opt: 256
 Smith-Waterman score: 408; 27.0% identity in 441 aa overlap

	10	20	30	40	50	
ma29gcf.pep	MGTDQGKT---FTWEELAHNTKDDLLLAI	RGRVYDVT	KFLSRHPGVDT	LLL	GAGRDT	
253538a	QGPTPRYFTWDEVAQRSGCEERWLVIDRKV	YNISEFTRRHPGGSRVISHYAGQDAT				
	10	20	30	40	50	
	60	70	80	90	100	110
ma29gcf.pep	PVFEMYHAF-GAADA	IMKKYYV	GTLVS	NELPIF	PEPTVFHK	TRVEGYFTDRNIDPKN
253538a	DPFVA	FHINKGLV	KYMN	SLLIGEL-SPEQPSF	EPTK	NKELTDEFREL
	60	70	80	90	100	110
	120	130	140	150	160	170
ma29gcf.pep	RPEIWGRYALIFG	SLIASYYAQLF	VPFV	VERTWLQVVF	-AIIMGF	ACAQVGLNPLHDASH
253538a	ANHVF--FLLYLLH	ILLDGA	AWTLWV	FGTSFLP	FLLCAV	LLSAVQAAQAGWLQ-HDYGH
	120	130	140	150	160	170
	180	190	200	210	220	
ma29gcf.pep	FSVTHNPTVW	KILGATHDF	---	FNGASYL	VWMYQHMLGH	HPTNIAGADPDVSTSE---
253538a	LSVYRKPK	-WNHL--VHKF	VIGHLKG	ASANWWNHRH	-FQHAKPNI	FHKDPDVNMLHV
	180	190	200	210	220	
	230	240	250	260	270	280
ma29gcf.pep	----PDVRR	IKPNQKWF	-VNHI	NQHMFV	--PFLYGL	AQVHHLFPNVS
253538a	LGEWQ	PIEYGKKL	KYLP	YNHQHEY	FLIGP	QI
	230	240	250	260	270	280
	290	300	310	320	330	340
ma29gcf.pep	NPISTWHTVM	FWGGKA	FVWYR	LIVPLQY	YPLGKV	LLFTVADMVSSYWL
253538a	----AWAVSYYI	--RFFITY	--IPF-YGILG	-ALLFLN	FIRFLE	SHWFVWVTQMNHIV
	290	300	310	320	330	340
	350	360	370	380	390	
ma29gcf.pep	EEVQWPLP	DENGIIQKDWAAMQVETT	---	QDYAHDS	HLWT	SITGSLNYQAVHH
253538a	MEI	----DQEAY	--RDWFSSQLT	ATCNVEQSFFND	--WFS	--GHLF
	340	350	360	370		
	400	410	420	430	440	
ma29gcf.pep	QHHYPDILAI	IKNTCSEY	KVPYLV	KDFTWQAFASH	LEHLRLV	GLRPKEEX
253538a	RHNLHKIA	PLVKS	CAKHGIEY	QEKP	LLRALLDI	IRSLKKSGKLWL
	380	390	400	410	420	430

Figure 9